DOCUMENT RESUME

ED 191 447 IR 008 679

AUTHOR Emerson, Jeff D.
TITLE The Use of On-Line Information Systems in Support of Complex Litigation.
PUB DATE Dec 79
NOTE 30p.: Paper presented at the National Online Information Meeting (New York, NY, March 25, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Court Litigation: Data Bases: Information Retrieval: *Information Systems: *Online Systems: Systems Development

ABSTRACT

This paper describes the characteristics of a computerized information retrieval system which could be used by attorneys engaged in such complex litigation as (1) cases which involve multi-district litigation: (2) cases involving complex issues of fact, as in regulatory agency actions: (3) cases involving a large number of documents: and (4) cases with a short preparation time. The system described is application- and document-oriented, stores both descriptive bibliographic data and full text information, and has a database capacity in excess of 100,000 document citations and an operational period of one to five years. The seven basic tasks involved in the creation and operation of the typical litigation management support system are detailed: system design, document control, document analysis, vocabulary control, data entry, database load, and search and retrieval processes. (FM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED191447

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE USE OF ON-LINE INFORMATION SYSTEMS IN SUPPORT OF COMPLEX LITIGATION

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jeff D. Emerson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

(c) JEFF D. EMERSON
December, 1979

I2008679

We live in a litigious society; not only is the number of litigations filed each year on the rise, but their level of complexity is also increasing. "In this revolutionary age, the rising flood of litigation in the courts of the United States and of the several states has nearly overwhelmed the Bench and Bar."[1] There are now close to one-half million lawyers in the United States. The practice of law is largely structured by tradition and custom. Relatively few attorneys have successfully embraced the technological advances which their sister professions, such as medicine, have come to rely upon. But this situation is changing. As Robert Arthur of Control Data Corporation has noted, "in a few years lawyers are going to exist in a world of digital information."[2] The reliance on computers to assist in the litigation process is inevitable, because the technology is both more readily available and because it is essential to the large-scale complex case.

CHARACTERISTICS OF A LITIGATION SUPPORT SYSTEM

A litigation support system which relies upon a computerized information retrieval system has become a virtual necessity for the successful litigation of a complex case. There are several characteristics of a complex litigation which may require an automated support system: those cases which involve multi-District litigation; cases involving complex issues of fact, as in regulatory agency actions; cases which involve a large number of documents; and cases with a short preparation time. Typically, cases which require an automated litigation support system are those designated as complex litigations. Within the last decade these litigations have dramatically increased in number and in their individual complexity. In the action brought by the Federal Trade Commission against the major oil companies, the discovery, "the pretrial

information-gathering process...has been contemplated at 40 million pages."[3] Faced with the requirement to acquire, collate, review, analyze, store and retrieve information from such an enormous collection of documents, the well-defined computer-aided litigation support system is indispensible.

The litigation support system is a specialized large-scale information system. It utilizes the same methods, tools, and techniques of other major information systems. In addition, most litigation support systems use a generalized automated information retrieval software package. The creation of the litigation data base is the crucial and distiguishing feature of the litigation support system. These systems usually share five charateristics:

(1) The litigation support system is **application oriented.** It is created as a customized information system for a specific litigation. In particular, the architecture of the data base is case-oriented in order to be responsive to the special needs of the attorneys. The data which are captured for an antitrust suit will likely be entirely different from those which are captured for a secuities fraud case.

(2) The data base is **document oriented.** Many large scale information systems are issue or fact-oriented. However, in ligigation systems, the correlation of a specific fact to a particular document is crucial. This is because the system is designed to correlate large amounts of information contained in specific documents. The documents are important for evidentiary submission. The retrieved information must be related to the physical document

from which the information was originally captured; and the original document must be addressed within the system. It should be possible to select one document page out of forty million.

(3) The litigation data base should support both **discrete, bibliographic data** which are descriptive of the documents and **full text information.** This flexibility of data representation is crucial to an adequate treatement of a wide variety in the relevance of documents in the system. The most highly relevant documents are typically captured in a full text format.

(4) The litigation support system data base is usually quite **large:** in excess of one hundred thousand (100,000) document citations. Several extant data bases contain as many as one-half million document citations.

(5) The data base which has been created for a specific litigation typically has a relatively **short lifespan.** The data base may be operational for a period of one to five years. Although some complex litigations have been pending for over ten years, the average lifespan of a litigation data base is probably no more than three years. Since the data base is created and maintained for a specific application, at the conclusion of the litigation the machine-readable data are archived or destroyed.

These characteristics of the litigation support system distinguish it from other types of large-scale information systems. However, the automated litigation support system does share more than methods and techniques with other large-scale information systems. It shares the information management efficiencies of computerized retrieval systems. Figure 1 is a graphic representation of the relationship between computer and manual based systems in the information management of large document collections. The efficiency level of information management in a very large document population, such as those encountered in litigation systems, drops off dramatically when the system is exclusively manual. The ability to efficiently manage the documents, collect and analyze relevant information, and substantiate issues are all critical needs of a litigation support system.

SYSTEM DESIGN

The computer is a relatively recent development in litigation support; it has been in active use only for the last five years. It is a critical aspect of a litigation support system, but by no means is the computer the driving force of the system. One of the worst mistakes in the design of any information system is to conform the needs of the users to the artificial requirements of a particular computer software system or data base architecture. The design of the litigation support system, particularly since it is so application oriented, is intended to be responsive to the needs of the attorneys. This design must necessarily be as flexible as possible because the nature of a complex litigation may change substantially as the case proceeds. Issues may change, new documents may be uncovered, threads of arguments by either or both sides may alter, or external events may cast a new complexion on the case.

mation Management
% EFFICIENCY
NUMBER OF DOCUMENTS

The design of the litigation support system, even more than with other large-scale information systems, must be carefully performed. The litigation support system is typically a hybrid of manual control of document processing and computer storage and retrieval of extracted information. The needs of both of these components, manual and computer, must be considered and accomodated. The manual component is concerned with the acquisition (discovery), screening, classification, review, analysis, encoding, and archiving of the physical documents. The computer component involves the data entry conversion of the encoded data into a machine-readable form, its load into a data base, and the subsequent retrieval of information from that data base. The processing of information in the litigation support system is essentially serial, with the true computer applications as the last stages in the processing cycle. The system often marries the data base to the physical document archive by addressing the physical document location or accessing microform images of the documents.

The acquisition of the documents is the first stage in the litigation support system processing cycle. Documents are typically acquired through the legal mechanism of discovery. Discovery may be an extended process, particularly in complex litigations.

> Discovery (is) designed to disclose (1) the names and location of witnesses whose written interrogatories or depositions upon oral interrogatories may be sought on the merits and (2) the existence, location and custody of documents and other physical evidence, the production of which may be sought on the merits.[4]

In a complex litigation, it is not unusual for the opposing party to innundate the discovering party with a hugh number of documents. This technique is designed to bury the discovering party with a mass of documentary materials, many of which may be only marginally relevant. Under such circumstances, the discovering party is under an even greater burden to control the documents for efficient processing into the system.

DOCUMENT CONTROL

Document control is essential for several reasons:

- It ensures the integrity of the document collection, some of which may later be required for submission into evidence.
- It ensures an unbroken chain of custody of the documents which is critical to evidentiary requirements.
- It prepares documents for entry into the litigation support system processing cycle.

There are four important components of the document control function: document numbering; the elimination of duplicates; the micrographics sub-system; and the reprographics sub-system.

Numbering is one of the most mechanically simple, but critical, tasks in the manual processing of the documents. The documents are numbered for several reasons: to catalog the amount and sequence of documents; to identify the source of the documents; to maintain the integrity of the paper file; and to provide a unique physical address for each document in the collection. This physical address is usually input to the data base for storage as part of the document citation. The documents are numbered prior to any other processing task. One important consideration in the numbering process is the ability to selectively mask the document numbers. This is done to preclude opposing counsel from inferring any identifying patterns or sequences in produced documents.

The elimination of duplicates in the document collection is a screening function which most often proves to be highly cost-effective. The single best way to reduce litigation support system costs is to reduce the volume of documents which are processed. The earlier that the document population can be reduced, the greater the cost reduction impact. The document populations of litigations typically have a high redundancy factor; occassionally as high as thirty to fifty percent (30-50%) of the documents are duplicated. In some measure, this redundancy factor is related to the type of litigation. An antitrust case may have a high redundancy factor, while for a securities case or an EEO case the redundancy factor may be quite low.

The elimination of duplicates in a litigation with a very large volume of uncollated documents can be very difficult. In general, a combination of manual and computer duplicate checking is the most effecient method of processing. Duplicate documents may be provisionally identified by an algorithm consisting of the document date, the author and the addressee. This minimal bibliographic information can be collected and compared to listings of document citations already on the data base. Potential duplicates can be identified and then researched to determine the degree of duplication. In a litigation, the existence of hand-written marginalia on otherwise duplicate documents can be sufficent to classify them as independent documents. The elimination of duplicate documents by clerical or para-legal personnel, prior to attorney review, can save substantial time for the litigating attorneys.

The micrographic sub-system is often an important component of the document control mechanism for a litigation support system. Microform systems are playing an increasing role in the support of computerized systems. There are several reasons for using microforms:

(1) The original documents, because they may have legal standing as evidence, are stored in an archive. Even the attorneys may have minimal access to the originals and then under very strict controls.

(2) The access to paper copies of the originals may also be restricted because:

- o the copies are used by the document catalogers and indexers the majority of the time; or
- o the copies take the place of the originals in the attorney work area and are therefore restricted.

(3) The use of any paper copy may be inadvisable because:

- o Transmitting even small quantities of paper from one location to another can be time-consuming and expensive.
- o Attorneys may conduct computer searches at a location where no paper copy is available, or at a time when access is not permitted, and therefore the paper copies would be useless.

In litigation support systems microfilm and microfiche have found wide acceptance because they overcome many of the drawbacks of paper:

- o Storage Space. Records on microfilm need only about four percent (4%) of the space required by paper copy. An entire document collection can be packed in a suitcase or briefcase and carried around the country.

- o File Integrity. Document images on microform are in a fixed sequence. There are no individual pages to become torn, lost, or misfiled. The potential for losing a microform record is minimal.

- o Paper Copy Availability. If a paper copy of a document is needed, the microform display equipment is capable of quickly providing a "blowback" of the microfilmed image.

- o Distribution. Microfilming provides a very convenient and economical method for providing duplicate sets of documents to multi-site locations.

- o Filing and Control. Filing and Control time are greatly reduced or eliminated altogether. The integrity of the document file is ensured.

- o Access Time. Microform, in connection with a computerized data base which provides the index to it, provides a rapid method for accessing images of the original documents.

There are several forms of microform that have been successfully used in conjunction with a computerized litigation support data base. The choice of microform is influenced by system characteristics and should be determined only after careful consideration of many factors, including filming costs, document preparation procedures, retreival requirements, equipment cost and availability.

Reprography is an optional feature in many litigation support systems. In general, the handling of original documents should be discouraged since they may be required for submission into evidence. In such cases, it is

advisable to create a working set of documents by xerographic reproduction. When there are several groups of counsel, and particularly when they are geographically separated, it is desirable to create a working set of documents for each group of attorneys as well as for the document coders. This can be a costly process: on one litigation the attorneys spent almost four million dollars on xerography alone.

The reproduction of several hundred thousand pages of documents requires the same controls and procedures as for micrographic processing. It is ironic that in an automated information system the most exposure is in the manual processing operations. During microfilming and reproduction, the documents have the highest risk of being damaged, mis-filed, or lost. A document numbering and control system is mandatory to ensure the integrity of the paper file.

The microform copy of the documents is generally kept as an archival file collection. The xerographic reproduction of the documents may be used by the attorneys as a working copy, or by the document analysts as a review copy.

DOCUMENT ANALYSIS

In order to extract the appropriate information from the documents for load into a computer data base, the documents must be reviewed and analyzed by document analysts. This document analysis function may involve a relatively simple extraction of bibliographic descriptive information from the face of the document; or it may involve a sophisticated interpretational indexing function.

A document may have objective data, keywords, an abstract or its full text recorded and loaded into the computer. The document analysis can be performed by any of the following methods, individually or in combination:

- o full text
- o extract
- o abstract
- o discrete (bibliographic) item
- o check-off

The full text approach involves the verbatim capture of the entire document. Full text is expensive because of the data entry and data storage costs. Because of these costs, it is usally reserved for the only most critical documents: complaints, briefs, depositions, and selected transcripts. Also, the use of full text data capture may have the legal consequence of precluding the document data base from the protection of attorney work product since full text documents may be considered as discoverable business records.

> It is interesting to note that if respondent employed a full text retrieval system and the opponent attempted to compel the respondent to use it to answer similar interrogatories, respondent would have had more difficulty arguing that opinions, theories, and mental impressions would be discovered. A full text system permits retrieval on the basis of any significant word or words in the full text of a document. Indeed, one strength of a full text system is that the selection of significant documents can be postponed until case issues are brought into sharp focus. Thus the system does not require the party to make any presupposition as to what the ultimate facts and issues will be, and the party may choose to make no significant screening of documents prior to introducing them into the data base. Even if the full text system were prepared in anticipation of litigation, however, the moving party would still have to show that it had a substantial need for the information and that it could not, without substantial hardship, obtain the materials by other means.(5)

An extract is a portion of the verbatim text of the document. It has the advantage of limiting the amount of information to only the most important and relevant portion(s) of the text. This reduces the cost without any loss of meaningful information. Moreover, this approach reflects a treatment of the document which requires a judgmental decision, thereby ensuring its consideration as attorney work product.

An abstract is a precis of the full text which is prepared by a document analyst. Abstracting is a difficult function to perform well since it requires a person who is knowledgeable in the subject matter of the documents as well as in abstracting principles.

A bibliographic item is a non-text datum, such as an author, a date, or a keyword. Bibliographic items are very helpful retrieval adjuncts to the text of a document.

A check-off item is a data field which can be defined by a yes/no, or present/not present decision. It may be used to indicate whether the document has been produced to opposing counsel.

These five methods are generally reduced in practice to three approaches to capture information from the documents: cataloging, indexing and abstracting.

Cataloging. The purpose of cataloging is to extract objective bibliographic information from the document. Examples of such information are names of authors and addressees of correspondence, their organ-

izational affiliation, titles or subjects of reports or memoranda, dates of publication or creation, page numbers, and the source of the document. These items supplement the text of the document.

Indexing. The indexing process extracts information, such as words or phrases from the text of the document, to capture the subject content. Typically, indexing involves the use of either a taxonomy or a thesaurus.

Abstracting. An abstract is an independently written summary which covers the main points of the document in a few short sentences. An extract is an excerpt of the document which can also be used to provide a summary of the contents of the document. An extract may be used when the document text is highly complex, or when the natural language of the text must be preserved for efficiency of retrieval.

The document analysis function is a critical aspect of the litigation support system. The quality of the information which is stored in the data base will be only as good as the quality of the data capture. One of the general ways to facilitate and control the document analysis operation is through the use of vocabulary control tools.

VOCABULARY CONTROL

A controlled vocabulary is often required as an adjunct to the process of creating a massive index to a large number of documents.

Two vocabulary control techniques which are commonly used are the taxonomy and the thesaurus. Both are carefully controlled vocabularies, embedding hier-

archical key term relationships, which greatly increases the precision of information retrieval. However, there are certain differences between the two techniques:

- o A taxonomy is a pre-determined vocabulary where the hierarchical relationships among terms are assigned prior to document coding. This approach decreases the time required to analyze, capture and input the document data, since a relatively few number of terms are normally involved.
- o A thesaurus is constructed both during and after the coding process. New terms are added as part of a vocabulary control process which is on-going during document review and thesaurus development. This approach provides a depth and flexibility which is not available with the taxonomy method. It does, however, increase the time to analyze, capture, and input the document data.

Vocabulary control is absolutely essential if the data base searcher is to have confidence in his results. The controlled vocabulary provides a common medium of communication between the document analyst and the data base searcher.

After the data capture is completed and the relevant information is recorded on coding worksheets, the data must be converted to a machine-readable format. This is the data entry function.

DATA ENTRY

There are a wide variety of data entry techniques, including:

- o Keypunching
- o Key to disc
- o Optical Character Recognition (OCR) typing
- o Automatic typing and word processing eqiupment
- o Stenographic court reporting conversion processes

The most commonly used data entry technique is still keypunching. This is also the least sophisticated and most error-prone method. The use of punch cards entails an additional problem in file integrity, since the individual cards may be mis-filed, damaged, or lost.

The most sophisticated data entry technique is key to disc operation. These systems generally operate under the control of a minicomputer. The computer may be programmed to provide forms display, tutorial prompts, and logical validations of the data. These techniques can save time and improve the accuracy of the data entry operator. However, key to disc systems are typically quite expensive to maintain and relatively difficult to learn. In addition, if the equipment fails, all of the data entry operators are inactive until it can be repaired.

Optical Character Recognition utilizes standard typewriters as the data entry mechanism. The typewriters are equipped with a special type element which can be read by an OCR scanner. The scanner converts the typed page to a machine-readable format. OCR has several advantages: it is easy to learn, the equipment investment is minimal (except for the scanner), and the downtime is low since the data entry mechanism is individual typewriters. The major disadvantage with OCR is that it usually requires a rigid format to be compatible with the scanner limitations.

Word processing equipment holds a promise for the future with the possibility that large numbers of documents in a typical office environment will initially be produced and stored in a machine-readable format. However, this potential is a long way off. Most litigation support systems make little use of this technology. Indeed, it is very unlikely that any substantial

number of litigation documents, other than case-specific briefs, will be available from word processing systems. Moreover, there are at present over one hundred different word processing systems. Only a small number of these systems utilize a common communications format. At least one large corporation uses over seventy different systems, none of which can talk to each other. Because of this problem, the practical use of word processing technology for the retrospective machine-readable capture of documents is limited.

The stenographic court reporting process is also limited in its practical application to litigation support systems. There are very few commercially available services. The few which are available use two basic techniques:

- A simultaneous capture of the stenographic shorthand in a machine-readable format. The shorthand is later translated, by computer, into English. Any translation errors are corrected by a human editor.
- The stenographic shorthand is translated by a human notereader who simultaneously transcribes the text into a word processing device.

The key point in both techniques is that the final transcript is produced from the machine-readable medium. It is essential that there is an identity between the printed and machine-readable copies of the transcript.

The stenographic court reporting process may be difficult to arrange, particularly if there are geographically distributed litigation sites. Moreover, while it is less expensive than re-keying the data, it typically costs at least $1.00 per page just for the machine-readable version.

Data entry remains one of the most difficult and error-prone functions in the litigation support system. The design and control of the data entry sub-system deserves careful attention. After the data have been converted into a machine-readable format, they are loaded into the data base.

DATA BASE LOAD

In the typical litigation support system, the cost of all of the computer operations is only about ten to fifteen percent (10-15%) of the total cost of the system. However, the computer operations are most often singled out for intense scrutiny, possibly because the information storage and retrieval functions are crucial to the efficiency and success of a litigation support system.

The data base load function should include a quality control check of all data which are updated to the data base. This may be accomplished by the use of an input pre-processor software package which performs edit checks on the input data, formats data base load transactions, and generates error listings and run statistic reports.

Each of the logical edit check validations may be specified in descriptive tables for the input pre-processor. These specifications ensure the consistency and the accuracy of the data loaded into the data base. Moreover, if they are externally defined by parameter, the software remains flexible for application-oriented edit routines. The input pre-processor should trap logical errors which are identified in the input record. Those records which contain detected errors can then be posted to an error suspense file which is disc resident. This holding file insures that the data fields which contain logical errors are not updated to the data base. The data fields on the error suspense

file can be corrected by an on-line computer terminal function, with the corrected error records being recycled through the input pre-processor to be re-validated and updated to the data base.

The input pre-processor software should be modular in design; its validation checks should be performed by externally defined parameters and specially designed application subroutines. It should be used for new data being loaded into the data base for the first time; corrected, recycled data which had errors detected in the first input run; and data base update maintenance transactions. Since the same program is used both to process new additions and update maintenance runs to the data base, a complete and timely audit trail report is insured.

In a litigation support system, updating the data base is essential. Three types of updates are necessary:

- Addition of new documents;
- Addition of new data to existing documents; and
- Correction of errors found on the data base.

Although some software systems allow on-line update of the data base, the best technique is batch processing update. There are several reasons for this assertion:

- Batch processing of update transactions is the least expensive and most efficient mechanism for the manipulation of data in an on-line data base. This is particularly true of large data base files. In addition, there is usually an extensive audit trail of transactions including the date, time, file, type of update transaction, the record updated, and the nature of the update.

- On-line interactive update of data base files can be extremely dangerous because:

  (a) The information in the data base is exposed to constant alteration, potentially even between two searches only minutes apart.

  (b) The audit trail for on-line update is usually confined to a hard copy terminal listing which results from the update session.

  (c) The integrity of the data base is open to violation by any user who has the appropriate access codes. Since the audit trail is also usually under the control of the user, it is theoretically possible in some systems to update the data base without the proper controls, approvals, or notifications.

- On-line interactive update of data base files does not allow for the edit checks on the update transactions which are normally performed by batch processing software.

A batch processing update technique, supported by an on-line text editor, combines the best of both update methods.

SEARCH AND RETRIEVAL

Once the data have been loaded to the data base, they are available for search and retrieval. There are two basic methods for information retrieval: batch and on-line. The batch retrieval system usually involves a significant delay between the search query and the system response. This is because the query is processed at a later time; the searcher does not have an immediate feedback from the system. In an on-line information retrieval system, the searcher conducts an interactive dialogue with the system; the results of a search query are immediately presented to the searcher, usually in less than ten seconds.

An on-line, inter-active system is critical to the success of information retrieval in a complex litigation support activity. There are several reasons for the requirement of an on-line system:

- The major portion of data base creation costs are in the document preparation, document analysis, data entry, and data base load operations. These operations are necessary preliminary steps to creating either an on-line or a batch retrieval system. The only real cost difference between on-line and batch retrieval systems is in the storage costs. On-line storage costs have been dropping steadily with the advancing hardware technology.

- A batch retrieval system is incapable of providing the immediacy, flexibility, and depth of response that an on-line interactive retrieval system can provide. An on-line interactive retrieval system allows the user to conduct a dialogue with the system. The results of a search request are immediately presented to the user for review and refinement.

- In the situation of a time-critical lawsuit, a batch job (turnaround for which is normally accomplished overnight) may not produce the results at a pace quick enough to satisfy the requirements of the attorneys. The opportunity cost of a fast response may be substantial.

- An on-line interactive system precludes the very real possibility of creating a search result which inadvertently prints five thousand pages of output from the data base. This is because the user, in a dialogue with the computer system, is in total control of the search session.

- The results of an on-line search are immediately known, informing the searcher not only of the answer to his question but also as to whether further research (expanding or contracting the inquiry) is warranted. An on-line system allows the user to refine his searches as the results become available. There is no lengthy wait time as there would be in a batch processing system.

A complex litigation support system demands the immediacy and flexibility of an on-line interactive retrieval system. This is unquestionably the most cost-efficient method of data base access.

An on-line interactive information retrieval system should provide certain key features in order to effectively support a complex litigation:

- The retrieval language should use a natural language command structure in a simple syntax. English-like commands are the most desirable. In addition, the retrieval system should be tutorial, so that unsophisticated users can learn the system during a search session.

- The retrieval system should allow for Boolean operation search arguments.

- The system should support vectored proximity search arguments. A vectored proximity search allows the system user to specify both the proximity and the order of occurrence of the text words. For example, a search argument WATER (3W) POLLUTION should locate text term occurrences of WATER POLLUTION and POLLUTION OF THE WATER.

- The retrieval system should allow special search capabilities, including root search; range search; proportional searching; and search masking. Proportional searching allows a user to enter a search argument which specifies a less than, greater than, or equal to relationship. Search masking allows the user to replace a character(s) in the search argument with a universal character which will match any legitimate character in an index term.

In addition to these retrieval oriented capabilities, the system should provide some basic output format displays. Ideally, the format display should be user-controlled for maximum flexibility in an application-oriented situation.

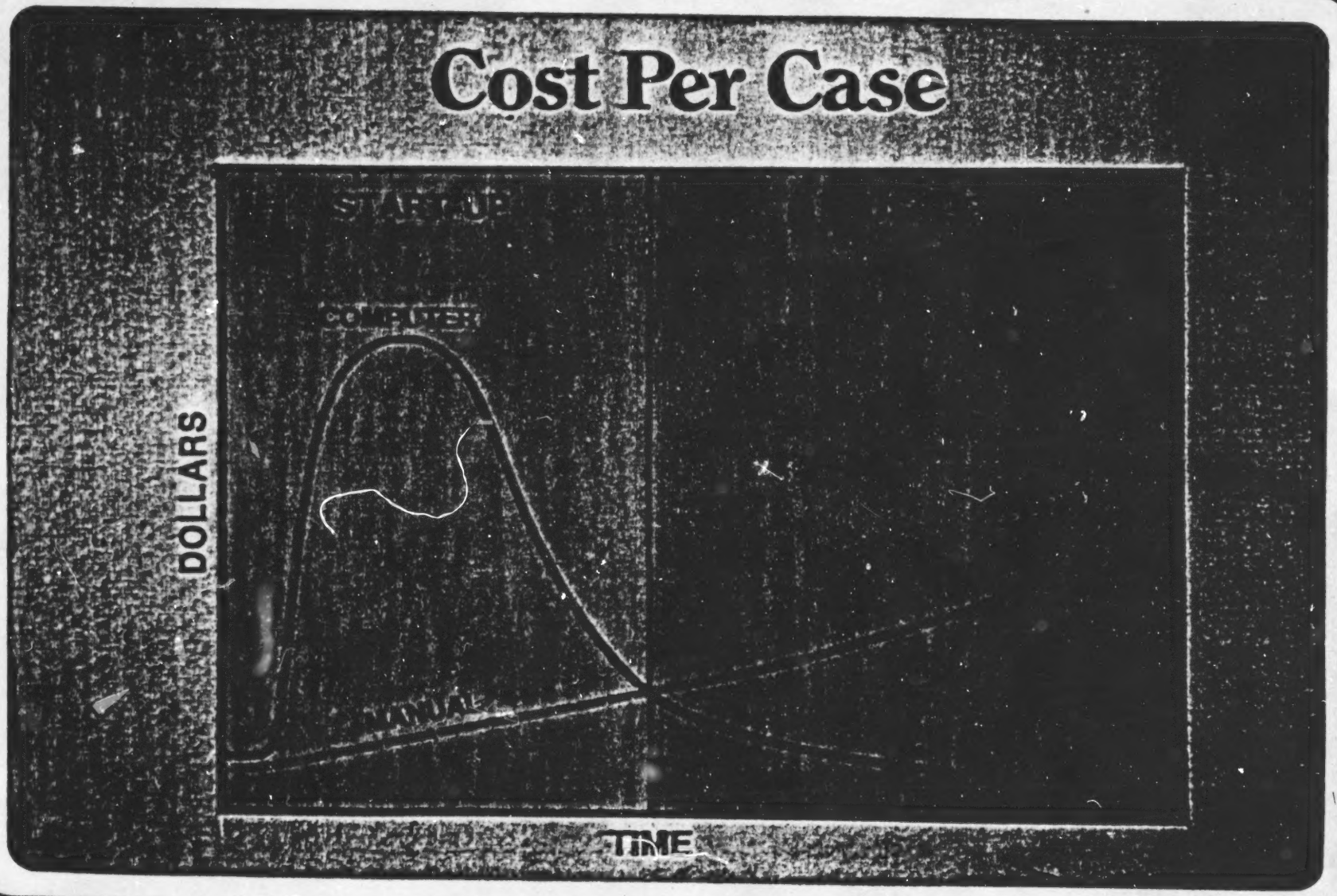
Most of the information retrieval in a litigation support system is oriented to queries for specific items of information: names, dates, and term references. The system does not need to be sophisticated in its search capabilities; but it does need to be fast to respond and easy to use. These are key points when the system users are highly paid attorneys who have limited information retrieval system experience.

COSTS

Litigation support systems can be very expensive. Because of this, not all cases warrant an automated support system. "The prevailing rule of thumb is that it takes 10,000 documents or more to justify the expense. A price tag of $400,000 to $600,000 for a computerized system isn't unusual and it can run into the seven figures."[6] One of the distinguishing features between a manual information system and a computerized information system is the start-up costs. These start-up costs are one-time system design and initiation expenses which must be amortized over the life of the system. In a computerized system, the start-up costs are disproportionate to the costs of maintaining the system. In a manual system, the start-up costs are typically low, but tend to increase in a linear fashion during the life of the system. This distinction is graphically depicted in the illustration in Figure 2.

The costs in a litigation support system are both volume and time sensitive. The only sure way to reduce the total system cost by an appreciable amount is to reduce the number of documents which are processed into the system. This can be achieved in several ways:

- The elimination of duplicate documents at an early stage of processing will reduce the total number of documents to be processed by one-tenth to as much as one-third.
- The documents should be treated in relation to their relevance to the litigation. High relevance documents are extensively analyzed; low relevance documents may have only a minimal amount of information recorded from them.
- The design architecture of the litigation support system data base should be simple, but flexible. If possible, the same design should be used to support several different cases, thereby allowing the design costs to be amortized over several litigations.


Cost Per Case
DOLLARS
TIME

- o Paralegal personnel should be involved in document screening, review, analysis, and retrieval. This allows the attorney to concentrate on case strategy and the actual litigation process.

- o The information retrieval system should be used as a shared resource, typically with access time purchased from a vendor. This reduces system purchase, equipment lease, and maintenance costs.

The average cost in a complex litigation support system is $6.00-$10.00 per document. Most of this cost is for the labor-intensive functions: design, document control, document analysis, lexicography, and data entry. In general the system design will account for about five percent of the total system cost; document control will account for about five percent; document analysis will be about sixty percent; data entry will be about ten percent; and the data base load will be about ten percent of the total system costs. The remaining ten percent will be spent in data base maintenance and retrieval functions.

It is important to recognize that the automated litigation support system, like any large-scale information mangement system, is still primarily labor-intensive. The largest proportion of system costs will be in the human labor functions. Because of this fact, variations in computer processing methods will likely have a small impact on the total cost of the system.

SUMMARY

The legal system is one of the last areas in our society to recognize and embrace the advantages of automated information management systems. With the dramatic increase in the number and complexity of litigations in recent years, some form of information management is essential. The litigation support system uses the techniques of information and computer science to control, correlate, store, and retrieve information in complex cases.

FOOTNOTES

(1) Manual for Complex Litigation. Chicago: Commerce Clearinghouse; 1978. p. xv

(2) as quoted by Salamon, Julie. "Computers are Gaining on the Front as Lawyers Aid in Complex Court Cases." The Wall Street Journal; November 2, 1977. p. 15

(3) Salamon; op. cit.

(4) Manual. op. cit. p. 3

(5) Fromholz, Haley J. "Discovery, Evidence, Confidentiality, and Security Problems Associated with the Use of Computer-Based Litigation Support Systems." Washington University Law Quarterly. Volume 1977 Number 3 Summer; p. 459

(6) Salamon; op. cit.